A semigroup method for high dimensional committor functions based on neural network

Haoya Li

Department of Mathematics, Stanford University

2nd MSML Conference

Joint work with Yuehaw Khoo, Yinuo Ren and Lexing Ying

Rare events encompass many phenomena in nature:

- chemical reactions
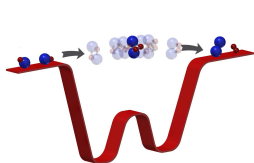
Rare events encompass many phenomena in nature:

- chemical reactions
- conformational change of biomolecules

Rare events encompass many phenomena in nature:

- chemical reactions
- conformational change of biomolecules
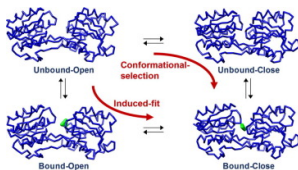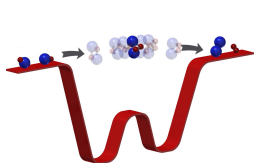

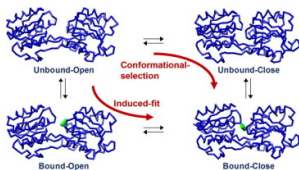(e) Chemical reaction

(f) Conformational change

Background – rare events

Rare events encompass many phenomena in nature:

- chemical reactions
- conformational change of biomolecules

(g) Chemical reaction   (h) Conformational change

The vibration of chemical bonds occurs on the time scale of 10^{-12} to 10^{-15} seconds, but a typical reaction may take seconds or longer to occur.

Intuitions of the system dynamics [1]

- A potential function is defined on the phase space of the system.

Intuitions of the system dynamics [1]

- A potential function is defined on the phase space of the system.
- The two metastable states (the reaction state A and the product state B) are two local minima of the potential function.

Intuitions of the system dynamics [1]

- A potential function is defined on the phase space of the system.
- The two metastable states (the reaction state A and the product state B) are two local minima of the potential function.
- There is some barrier between A and B.

Intuitions of the system dynamics [1]

- A potential function is defined on the phase space of the system.
- The two metastable states (the reaction state A and the product state B) are two local minima of the potential function.
- There is some barrier between A and B.
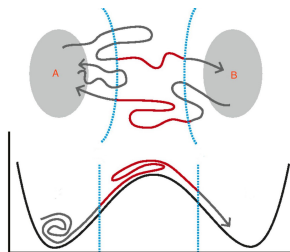- The system is subject to some noise, which leads to transitions from one state to another.
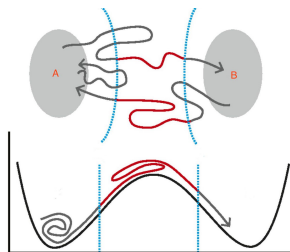
Intuitions of the system dynamics [1]

- A potential function is defined on the phase space of the system.
- The two metastable states (the reaction state A and the product state B) are two local minima of the potential function.
- There is some barrier between A and B.
- The system is subject to some noise, which leads to transitions from one state to another.
- The transition is rare since the noise is relatively small compared to the barrier.

How to model the dynamics between A and B?
transition mechanism, reaction rate, etc.

- Transition state theory (TST)
  In order to go from A to B, the system has to go to a saddle point on the potential energy landscape

How to model the dynamics between A and B?
transition mechanism, reaction rate, etc.

- Transition state theory (TST)
  In order to go from A to B, the system has to go to a saddle point on the potential energy landscape

- Large deviation theory (LDT)
  With LDT, you can calculate the probability that a diffusion process stays in a neighborhood of a particular path. The most probable transition path can be defined.

Limitations of TST and LDT

- In TST, it is assumed that every crossing of the transition state leads to a transition, which may overestimate the reaction rate.

Limitations of TST and LDT

- In TST, it is assumed that every crossing of the transition state leads to a transition, which may overestimate the reaction rate.
- The barriers may be entropic (especially when the dimension of the phase space is high), so the saddle point may not play the role of a transition state.

Limitations of TST and LDT

- In TST, it is assumed that every crossing of the transition state leads to a transition, which may overestimate the reaction rate.
- The barriers may be entropic (especially when the dimension of the phase space is high), so the saddle point may not play the role of a transition state.
- There can be an ensemble of paths contributing to the transition (instead of the most probable path itself).

System dynamics:

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right)dt + \sqrt{2\beta^{-1}}d\mathbf{w}_t \quad \text{(over-damped langevin)}$$

System dynamics:

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t \quad \text{(over-damped langevin)}$$

Backward operator:

$$L = \beta^{-1}\Delta - \nabla V \cdot \nabla$$

Problem setting – Transition path theory (TPT)

System dynamics:

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t \quad \text{(over-damped langevin)}$$

Backward operator:

$$L = \beta^{-1}\Delta - \nabla V \cdot \nabla$$

Forward operator:

$$L^*\phi = \beta^{-1}\Delta\phi + \nabla \cdot (V\phi)$$

Problem setting – Transition path theory (TPT)

System dynamics:

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t \quad \text{(over-damped langevin)}$$

Backward operator:

$$L = \beta^{-1}\Delta - \nabla V \cdot \nabla$$

Forward operator:

$$L^*\phi = \beta^{-1}\Delta\phi + \nabla \cdot (V\phi)$$

Equilibrium distribution:

$$\rho(\mathbf{x}) = \frac{1}{Z_\beta}\exp(-\beta V(\mathbf{x})) \quad (L^*\rho = 0)$$

Committor function:

$$q(\mathbf{x}) = \mathbb{P}(\tau_B < \tau_A \mid \mathbf{x}_0 = \mathbf{x}),$$

where τ_A and τ_B are the hitting times for the sets A and B, respectively.

Committor function:

$$q(\mathbf{x}) = \mathbb{P}(\tau_B < \tau_A \mid \mathbf{x}_0 = \mathbf{x}),$$

where τ_A and τ_B are the hitting times for the sets A and B, respectively.

Fokker-Planck (backward Kolmogorov) equation:

$$(-1/\beta\Delta + \nabla V \cdot \nabla)q = 0 \text{ in } \Omega\backslash(A \cup B), \quad q(\mathbf{x})|_{\partial A} = 0, \quad q(\mathbf{x})|_{\partial B} = 1.$$

Why is the committor function important?

- Transition rate:

$$\nu_R = \beta^{-1} \int_\Omega |\nabla q(\mathbf{x})|^2 \rho(\mathbf{x}) \mathrm{d}\mathbf{x}$$

Why is the committor function important?

- Transition rate:

$$\nu_R = \beta^{-1} \int_\Omega |\nabla q(\mathbf{x})|^2 \rho(\mathbf{x}) \mathrm{d}\mathbf{x}$$

- Probability density of reactive trajectories:

$$\rho_R(\mathbf{x}) \propto q(\mathbf{x})(1 - q(\mathbf{x}))\rho(\mathbf{x})$$

Why is the committor function important?

- Transition rate:
$$\nu_R = \beta^{-1} \int_\Omega |\nabla q(\mathbf{x})|^2 \rho(\mathbf{x}) \mathrm{d}\mathbf{x}$$

- Probability density of reactive trajectories:
$$\rho_R(\mathbf{x}) \propto q(\mathbf{x})(1 - q(\mathbf{x}))\rho(\mathbf{x})$$

- Reactive current
$$J_R(\mathbf{x}) = \beta^{-1} \rho(\mathbf{x}) \nabla q(\mathbf{x})$$

- **High-dimensionality of phase space**
  Classical methods such as finite difference, finite element methods suffer from the curse of dimensionality.

- **High-dimensionality of phase space**
  Classical methods such as finite difference, finite element methods suffer from the curse of dimensionality.

- **Singularity of the committor function**

  In the high T regime, the committor function can be steep near A, B since the FP equation converges heuristically to a Laplace equation.

  In the low T regime, there is typically a sharp interface between A and B.

- **High-dimensionality of phase space**
  Classical methods such as finite difference, finite element methods suffer from the curse of dimensionality.

- **Singularity of the committor function**

  In the high T regime, the committor function can be steep near A, B since the FP equation converges heuristically to a Laplace equation.

  In the low T regime, there is typically a sharp interface between A and B.

- **Enforcement of the boundary conditions**
  Large penalty terms may lead to a ill-conditioned problem, and reparametrization results in a much more complicated equation.

- **High-dimensionality of phase space**
  Parameterize the committor function with a neural network, and solve the optimization problem derived from the variational form.

Solutions in previous work [2]

- **High-dimensionality of phase space**
  Parameterize the committor function with a neural network, and solve the optimization problem derived from the variational form.

- **Singularity of the committor function**
  Use $\tanh$ activation for the last layer of the neural network, and add explicit singularity terms

Solutions in previous work [2]

- **High-dimensionality of phase space**
  Parameterize the committor function with a neural network, and solve the optimization problem derived from the variational form.

- **Singularity of the committor function**
  Use $\tanh$ activation for the last layer of the neural network, and add explicit singularity terms

  $$q_\theta(\mathbf{x}) := n_{\theta_A}(\mathbf{x})S_A\left(\mathbf{x} - \mathbf{y}^A\right) + n_{\theta_B}(\mathbf{x})S_B\left(\mathbf{x} - \mathbf{y}^B\right) + n_{\theta_0}(\mathbf{x})$$

Solutions in previous work [2]

- **High-dimensionality of phase space**
  Parameterize the committor function with a neural network, and solve the optimization problem derived from the variational form.

- **Singularity of the committor function**
  Use tanh activation for the last layer of the neural network, and add explicit singularity terms

$$q_\theta(\mathbf{x}) := n_{\theta_A}(\mathbf{x})S_A\left(\mathbf{x} - \mathbf{y}^A\right) + n_{\theta_B}(\mathbf{x})S_B\left(\mathbf{x} - \mathbf{y}^B\right) + n_{\theta_0}(\mathbf{x})$$
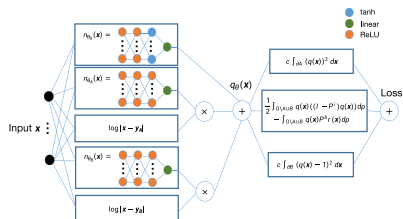
Solutions in previous work [2]

- **High-dimensionality of phase space**
  Parameterize the committor function with a neural network, and solve the optimization problem derived from the variational form.

- **Singularity of the committor function**
  Use `tanh` activation for the last layer of the neural network, and add explicit singularity terms

$$q_\theta(\mathbf{x}) := n_{\theta_A}(\mathbf{x}) S_A \left(\mathbf{x} - \mathbf{y}^A \right) + n_{\theta_B}(\mathbf{x}) S_B \left(\mathbf{x} - \mathbf{y}^B \right) + n_{\theta_0}(\mathbf{x})$$

- Enforcement of the boundary conditions

Solutions in previous work [2]

Instead of the strong form:

$$(-1/\beta\Delta + \nabla V \cdot \nabla)q = 0 \text{ in } \Omega\backslash(A \cup B), \quad q(\mathbf{x})|_{\partial A} = 0, \quad q(\mathbf{x})|_{\partial B} = 1,$$

we can work on the variational form:

$$\operatorname*{argmin}_{q} \int_{\Omega\backslash(A\cup B)} |\nabla q(\mathbf{x})|^2 \rho(\mathbf{x})d\mathbf{x}, \quad q(\mathbf{x})|_{\partial A} = 0, \quad q(\mathbf{x})|_{\partial B} = 1.$$

In [2], the boundary conditions are enforced through penalty functions

$$\operatorname*{argmin}_{\theta\in\mathbb{R}^p} \int_{\Omega\backslash(A \cup B)} |\nabla q_\theta(\mathbf{x})|^2 \rho(\mathbf{x})d\mathbf{x} + \tilde{c}\int_{\partial A} q_\theta(\mathbf{x})^2 dm_{\partial A}(\mathbf{x}) + \tilde{c}\int_{\partial B} (q_\theta(\mathbf{x}) - 1)^2 dm_{\partial B}(\mathbf{x}),$$

In the existing method,

- explicit differentiation is needed in the objective function, and
- the penalty coefficient needs to be carefully tuned since the boundary condition is enforced solely through the penalty terms.

In the existing method,

- explicit differentiation is needed in the objective function, and
- the penalty coefficient needs to be carefully tuned since the boundary condition is enforced solely through the penalty terms.

Intuitively, we can consider

$$(I - e^{\delta L})q = 0,$$

instead of

$$Lq = 0,$$

then the explicit differentiation is removed and Monte Carlo methods can be applied to simulate the semigroup operator $e^{\delta L}$.

Justification

Consider the Langevin process starting from a point $\mathbf{x} \in \Omega$

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t,$$
$$\mathbf{x}_0 = \mathbf{x}. \tag{1}$$

Justification

Consider the Langevin process starting from a point $\mathbf{x} \in \Omega$

$$
\begin{aligned}
d\mathbf{x}_t &= -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t, \\
\mathbf{x}_0 &= \mathbf{x}.
\end{aligned}
\tag{1}
$$

Proposition 1

When ∇V is bounded and Lipschitz continuous on $\mathbb{R}^d$, the committor function q satisfies the following semigroup formulation:

$$
q(\mathbf{x}) = (Pq)(\mathbf{x}) \quad in \quad \Omega \backslash (A \cup B), \quad q|_{\partial A} = 0, \quad q|_{\partial B} = 1,
\tag{2}
$$

Justification

Consider the Langevin process starting from a point $\mathbf{x} \in \Omega$

$$
\begin{aligned}
d\mathbf{x}_t &= -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t, \\
\mathbf{x}_0 &= \mathbf{x}.
\end{aligned}
\tag{1}
$$

Proposition 1

When ∇V is bounded and Lipschitz continuous on $\mathbb{R}^d$, the committor function q satisfies the following semigroup formulation:

$$
q(\mathbf{x}) = (Pq)(\mathbf{x}) \quad in \quad \Omega \backslash (A \cup B), \quad q|_{\partial A} = 0, \quad q|_{\partial B} = 1,
\tag{2}
$$

where the semigroup operator P is defined as

$$
(Pf)(\mathbf{x}) := \mathbb{E}^{\mathbf{x}}\left(f\left(\mathbf{x}_{\tau \wedge \delta}\right)\right),
\tag{3}
$$

where $\mathbb{E}^{\mathbf{x}}$ is the expectation taken with respect to the law of the process (1) and $\tau = \tau_A \wedge \tau_B$.

Justification

Consider the Langevin process starting from a point $\mathbf{x} \in \Omega$

$$d\mathbf{x}_t = -\nabla V\left(\mathbf{x}_t\right) dt + \sqrt{2\beta^{-1}} d\mathbf{w}_t, \tag{1}$$
$$\mathbf{x}_0 = \mathbf{x}.$$

Proposition 1

When ∇V is bounded and Lipschitz continuous on $\mathbb{R}^d$, the committor function q satisfies the following semigroup formulation:

$$q(\mathbf{x}) = (Pq)(\mathbf{x}) \quad in \quad \Omega\backslash(A \cup B), \quad q|_{\partial A} = 0, \quad q|_{\partial B} = 1, \tag{2}$$

where the semigroup operator P is defined as

$$(Pf)(\mathbf{x}) := \mathbb{E}^{\mathbf{x}}\left(f\left(\mathbf{x}_{\tau \wedge \delta}\right)\right), \tag{3}$$

where $\mathbb{E}^{\mathbf{x}}$ is the expectation taken with respect to the law of the process (1) and $\tau = \tau_A \wedge \tau_B$.

Proof.

Dynkin's formula [4]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Resolving the boundary conditions

We can split the semigroup operator P into two parts.

$$(Pq)(\mathbf{x}) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\right) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right)$$

Resolving the boundary conditions

We can split the semigroup operator P into two parts.

$$(Pq)(\mathbf{x}) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\right) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right)$$
$$= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right)$$

Resolving the boundary conditions

We can split the semigroup operator P into two parts.

$$
\begin{aligned}
(Pq)(\mathbf{x}) &= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau\wedge\delta}\right)\right) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau\wedge\delta}\right)\mathbf{1}_{\{\delta<\tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau\wedge\delta}\right)\mathbf{1}_{\{\delta\geq\tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right)\mathbf{1}_{\{\delta<\tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau}\right)\mathbf{1}_{\{\delta\geq\tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right)\mathbf{1}_{\{\delta<\tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(r\left(\mathbf{x}_{\tau}\right)\mathbf{1}_{\{\delta\geq\tau\}}\right) \\
&= (P^{i}q)(\mathbf{x}) + (P^{b}r)(\mathbf{x}),
\end{aligned}
\tag{4}
$$

Resolving the boundary conditions

We can split the semigroup operator P into two parts.

$$
\begin{aligned}
(Pq)(\mathbf{x}) &= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\right) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right)\mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(r\left(\mathbf{x}_{\tau}\right)\mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= (P^i q)(\mathbf{x}) + (P^b r)(\mathbf{x}),
\end{aligned} \tag{4}
$$

where r is a function defined on $\partial A \cup \partial B$ with $r(\mathbf{x})|_{\partial A} = 0$ and $r(\mathbf{x})|_{\partial B} = 1$.

Resolving the boundary conditions

We can split the semigroup operator P into two parts.

$$
\begin{aligned}
(Pq)(\mathbf{x}) &= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right)\right) = \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right) \mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau \wedge \delta}\right) \mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right) \mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\tau}\right) \mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= \mathbb{E}^{\mathbf{x}}\left(q\left(\mathbf{x}_{\delta}\right) \mathbf{1}_{\{\delta < \tau\}}\right) + \mathbb{E}^{\mathbf{x}}\left(r\left(\mathbf{x}_{\tau}\right) \mathbf{1}_{\{\delta \geq \tau\}}\right) \\
&= (P^{i}q)(\mathbf{x}) + (P^{b}r)(\mathbf{x}),
\end{aligned}
\tag{4}
$$

where r is a function defined on $\partial A \cup \partial B$ with $r(\mathbf{x})|_{\partial A} = 0$ and $r(\mathbf{x})|_{\partial B} = 1$.

The equation (2) can then be expressed by

$$
(I - P^{i})q(\mathbf{x}) - (P^{b}r)(\mathbf{x}) = 0.
$$

Note that the boundary condition is naturally included in the P^{b} term.

We have the follwing property on the positive definiteness of P^i:

We have the follwing property on the positive definiteness of P^i:

Proposition 2

P^i *is a symmetric operator on* $L^2_\rho(\Omega\backslash(A\cup B))$, *in other words,*
$\langle u, P^i v\rangle_\rho = \langle P^i u, v\rangle_\rho$, *where* $\langle f, g\rangle_\rho = \int_{\Omega\backslash(A\cup B)} f(\mathbf{x})g(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$ *is the inner product of the Hilbert space* $L^2_\rho(\Omega\backslash(A\cup B))$.

We have the follwing property on the positive definiteness of P^i:

Proposition 2

P^i is a symmetric operator on $L^2_\rho(\Omega\backslash(A\cup B))$, in other words,
$\langle u, P^i v\rangle_\rho = \langle P^i u, v\rangle_\rho$, where $\langle f, g\rangle_\rho = \int_{\Omega\backslash(A\cup B)} f(\mathbf{x})g(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$ is the inner product of the Hilbert space $L^2_\rho(\Omega\backslash(A\cup B))$.

Based on Proposition 2, we can derive the following variational formulation:

Positive definiteness and the new variational form

We have the follwing property on the positive definiteness of P^i:

Proposition 2

P^i is a symmetric operator on $L^2_\rho(\Omega\backslash(A\cup B))$, in other words,
$\langle u, P^i v\rangle_\rho = \langle P^i u, v\rangle_\rho$, where $\langle f, g\rangle_\rho = \int_{\Omega\backslash(A\cup B)} f(\mathbf{x})g(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$ is the inner
product of the Hilbert space $L^2_\rho(\Omega\backslash(A\cup B))$.

Based on Proposition 2, we can derive the following variational formulation:

$$\min_q \frac{1}{2}\int_{\Omega\backslash(A\,\cup\,B)} q(\mathbf{x})\Big((I-P^i)q(\mathbf{x})\Big)\rho(\mathbf{x})d\mathbf{x} - \int_{\Omega\backslash(A\,\cup\,B)} q(\mathbf{x})P^b r(\mathbf{x})\rho(\mathbf{x})d\mathbf{x} \quad (5)$$
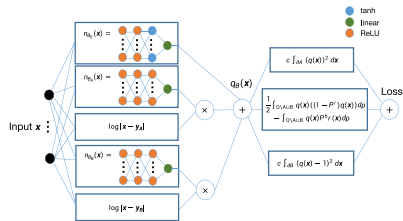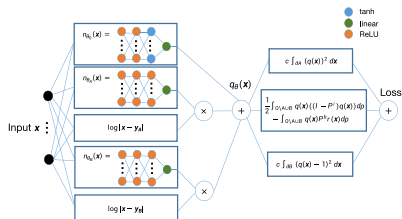
Nonlinear parametrization

We parametrize the committor function by an NN q_θ,

Nonlinear parametrization
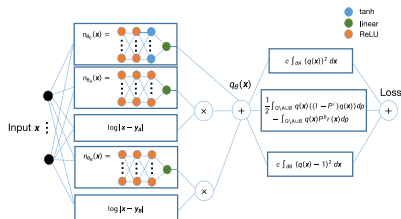
We parametrize the committor function by an NN q_θ,

Nonlinear parametrization

We parametrize the committor function by an NN q_θ,

then the optimization problem in terms of θ is

$$\min_\theta \frac{1}{2} \int_{\Omega \setminus (A \cup B)} q_\theta(\mathbf{x}) \left((I - P^i) q_\theta(\mathbf{x}) \right) \rho(\mathbf{x}) d\mathbf{x} - \int_{\Omega \setminus (A \cup B)} q_\theta(\mathbf{x}) P^b r(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}$$
$$+ \frac{c}{2} \int q_\theta(\mathbf{x})^2 dm_A(\mathbf{x}) + \frac{c}{2} \int (q_\theta(\mathbf{x}) - 1)^2 dm_B(\mathbf{x}), \tag{6}$$

Nonlinear parametrization

We parametrize the committor function by an NN q_θ,

then the optimization problem in terms of θ is

$$\min_\theta \frac{1}{2} \int_{\Omega\setminus(A\,\cup\,B)} q_\theta(\mathbf{x})\left((I - P^i)q_\theta(\mathbf{x})\right)\rho(\mathbf{x})d\mathbf{x} - \int_{\Omega\setminus(A\,\cup\,B)} q_\theta(\mathbf{x})P^b r(\mathbf{x})\rho(\mathbf{x})d\mathbf{x}$$
$$+ \frac{c}{2} \int q_\theta(\mathbf{x})^2 dm_A(\mathbf{x}) + \frac{c}{2} \int (q_\theta(\mathbf{x}) - 1)^2 dm_B(\mathbf{x}), \tag{6}$$

where we include the penalty terms to achieve a better numerical performance. We will show in the numerical experiments that the solutions obtained are **not** sensitive to the penalty coefficient.

In the application of SGD type methods to the optimization problem (6), we need to calculate the derivative for each of the terms.

Derivatives

In the application of SGD type methods to the optimization problem (6), we need to calculate the derivative for each of the terms.

- The first two terms (using the symmetry stated in Proposition 2):

$$
\int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left((I - P^i) q_\theta(\mathbf{x}) \right) \rho(\mathbf{x}) d\mathbf{x} - \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) P^b r(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}
$$

$$
= \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right) \rho(\mathbf{x}) d\mathbf{x}
$$

$$
= \mathbb{E}_{\mathbf{x} \sim \rho} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right),
\tag{7}
$$

Derivatives

In the application of SGD type methods to the optimization problem (6), we need to calculate the derivative for each of the terms.

- The first two terms (using the symmetry stated in Proposition 2):

$$\int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left((I - P^i) q_\theta(\mathbf{x}) \right) \rho(\mathbf{x}) d\mathbf{x} - \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) P^b r(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}$$

$$= \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right) \rho(\mathbf{x}) d\mathbf{x}$$

$$= \mathbb{E}_{\mathbf{x} \sim \rho} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right), \tag{7}$$

so an unbiased estimator is

$$\nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} - r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right). \tag{8}$$

Derivatives

In the application of SGD type methods to the optimization problem (6), we need to calculate the derivative for each of the terms.

- The first two terms (using the symmetry stated in Proposition 2):

$$\int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left((I - P^i) q_\theta(\mathbf{x}) \right) \rho(\mathbf{x}) d\mathbf{x} - \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) P^b r(\mathbf{x}) \rho(\mathbf{x}) d\mathbf{x}$$

$$= \int_{\Omega \setminus (A \cup B)} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right) \rho(\mathbf{x}) d\mathbf{x}$$

$$= \mathbb{E}_{\mathbf{x} \sim \rho} \nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - \mathbb{E}^{\mathbf{x}} \left(q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} \right) - \mathbb{E}^{\mathbf{x}} \left(r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right) \right),$$
(7)

  so an unbiased estimator is

$$\nabla_\theta q_\theta(\mathbf{x}) \left(q_\theta(\mathbf{x}) - q_\theta(\mathbf{x}_\delta) \mathbf{1}_{\{\delta < \tau\}} - r(\mathbf{x}_\tau) \mathbf{1}_{\{\delta \geq \tau\}} \right).$$
(8)

- For the third and fourth terms, unbiased estimators of their gradients are

$$c \nabla_\theta q_\theta(\mathbf{x}_A) q_\theta(\mathbf{x}_A), \quad c \nabla_\theta q_\theta(\mathbf{x}_B) (q_\theta(\mathbf{x}_B) - 1),$$
(9)

  respectively, where $\mathbf{x}_A \sim m_A$ and $\mathbf{x}_B \sim m_B$.

Monte Carlo sampling

In order to obtain the unbiased estimators above, we need to give samples for $\mathbf{x} \sim \rho$, the corresponding $\mathbf{x}_\delta$, and the indicators $\mathbf{1}_{\{\delta < \tau\}}$, $\mathbf{1}_{\{\delta \geq \tau = \tau_A\}}$ and $\mathbf{1}_{\{\delta \geq \tau = \tau_B\}}$.

- $\mathbf{x} \sim \rho$:
  In this paper, the potential is assumed to be confining, so the Langevin dynamics is ergodic, and $\mathbf{x} \sim \rho$ can be approximated by $\tilde{\mathbf{x}}_{N\Delta t}$ for a sufficiently small Δt and sufficiently large N with an arbitrary $\tilde{\mathbf{x}}_0$, where $\tilde{\mathbf{x}}_{N\Delta t}$ is obtained by the Euler-Maruyama scheme:

$$\tilde{\mathbf{x}}_{(n+1)\Delta t} = \tilde{\mathbf{x}}_{n\Delta t} - \nabla V(\tilde{\mathbf{x}}_{n\Delta t})\Delta t + \sqrt{2\beta^{-1}}\mathbf{w}_{\Delta t}.$$

Monte Carlo sampling

In order to obtain the unbiased estimators above, we need to give samples for $\mathbf{x} \sim \rho$, the corresponding $\mathbf{x}_\delta$, and the indicators $\mathbf{1}_{\{\delta < \tau\}}$, $\mathbf{1}_{\{\delta \geq \tau = \tau_A\}}$ and $\mathbf{1}_{\{\delta \geq \tau = \tau_B\}}$.

- $\mathbf{x} \sim \rho$:
  In this paper, the potential is assumed to be confining, so the Langevin dynamics is ergodic, and $\mathbf{x} \sim \rho$ can be approximated by $\tilde{\mathbf{x}}_{N\Delta t}$ for a sufficiently small Δt and sufficiently large N with an arbitrary $\tilde{\mathbf{x}}_0$, where $\tilde{\mathbf{x}}_{N\Delta t}$ is obtained by the Euler-Maruyama scheme:

$$\tilde{\mathbf{x}}_{(n+1)\Delta t} = \tilde{\mathbf{x}}_{n\Delta t} - \nabla V(\tilde{\mathbf{x}}_{n\Delta t})\Delta t + \sqrt{2\beta^{-1}}\mathbf{w}_{\Delta t}.$$

- $\mathbf{x}_\delta$
  Given $\mathbf{x} \sim \rho$, we approximate $\mathbf{x}_\delta$ by Euler-Maruyama scheme as well:

$$\mathbf{x}_\delta = \mathbf{x} - \nabla V(\mathbf{x})\delta + \sqrt{2\beta^{-1}}\mathbf{w}_\delta. \tag{10}$$

Monte Carlo sampling

In order to obtain the unbiased estimators above, we need to give samples for $\mathbf{x} \sim \rho$, the corresponding $\mathbf{x}_\delta$, and the indicators $\mathbf{1}_{\{\delta < \tau\}}$, $\mathbf{1}_{\{\delta \geq \tau = \tau_A\}}$ and $\mathbf{1}_{\{\delta \geq \tau = \tau_B\}}$.

- $\mathbf{x} \sim \rho$:
  In this paper, the potential is assumed to be confining, so the Langevin dynamics is ergodic, and $\mathbf{x} \sim \rho$ can be approximated by $\tilde{\mathbf{x}}_{N\Delta t}$ for a sufficiently small Δt and sufficiently large N with an arbitrary $\tilde{\mathbf{x}}_0$, where $\tilde{\mathbf{x}}_{N\Delta t}$ is obtained by the Euler-Maruyama scheme:

$$\tilde{\mathbf{x}}_{(n+1)\Delta t} = \tilde{\mathbf{x}}_{n\Delta t} - \nabla V(\tilde{\mathbf{x}}_{n\Delta t})\Delta t + \sqrt{2\beta^{-1}}\mathbf{w}_{\Delta t}.$$

- $\mathbf{x}_\delta$
  Given $\mathbf{x} \sim \rho$, we approximate $\mathbf{x}_\delta$ by Euler-Maruyama scheme as well:

$$\mathbf{x}_\delta = \mathbf{x} - \nabla V(\mathbf{x})\delta + \sqrt{2\beta^{-1}}\mathbf{w}_\delta. \tag{10}$$

- Indicators:
  The following approximations are used for the indicators:

$$\mathbf{1}_{\{\delta < \tau\}} = 1 \text{ if } \mathbf{x}_\delta \in \Omega \backslash A \cup B,$$
$$\mathbf{1}_{\{\delta \geq \tau = \tau_A\}} = 1 \text{ if } \mathbf{x}_\delta \in A, \quad \mathbf{1}_{\{\delta \geq \tau = \tau_B\}} = 1 \text{ if } \mathbf{x}_\delta \in B.$$

Monte Carlo sampling

In order to obtain the unbiased estimators above, we need to give samples for $\mathbf{x} \sim \rho$, the corresponding $\mathbf{x}_\delta$, and the indicators $\mathbf{1}_{\{\delta < \tau\}}$, $\mathbf{1}_{\{\delta \geq \tau = \tau_A\}}$ and $\mathbf{1}_{\{\delta \geq \tau = \tau_B\}}$.

- $\mathbf{x} \sim \rho$:
  In this paper, the potential is assumed to be confining, so the Langevin dynamics is ergodic, and $\mathbf{x} \sim \rho$ can be approximated by $\tilde{\mathbf{x}}_{N\Delta t}$ for a sufficiently small Δt and sufficiently large N with an arbitrary $\tilde{\mathbf{x}}_0$, where $\tilde{\mathbf{x}}_{N\Delta t}$ is obtained by the Euler-Maruyama scheme:
  $$\tilde{\mathbf{x}}_{(n+1)\Delta t} = \tilde{\mathbf{x}}_{n\Delta t} - \nabla V(\tilde{\mathbf{x}}_{n\Delta t})\Delta t + \sqrt{2\beta^{-1}}\mathbf{w}_{\Delta t}.$$

- $\mathbf{x}_\delta$
  Given $\mathbf{x} \sim \rho$, we approximate $\mathbf{x}_\delta$ by Euler-Maruyama scheme as well:
  $$\mathbf{x}_\delta = \mathbf{x} - \nabla V(\mathbf{x})\delta + \sqrt{2\beta^{-1}}\mathbf{w}_\delta. \tag{10}$$

- Indicators:
  The following approximations are used for the indicators:
  $$\mathbf{1}_{\{\delta < \tau\}} = 1 \text{ if } \mathbf{x}_\delta \in \Omega \backslash A \cup B,$$
  $$\mathbf{1}_{\{\delta \geq \tau = \tau_A\}} = 1 \text{ if } \mathbf{x}_\delta \in A, \quad \mathbf{1}_{\{\delta \geq \tau = \tau_B\}} = 1 \text{ if } \mathbf{x}_\delta \in B.$$

- Multi-step Euler-Maruyama can be adopted to improve accuracy.

Numerical experiment I – Double well potential

The double well potential we consider is given by:

$$V(\mathbf{x}) = \left(x_1^2 - 1\right)^2 + 0.3 \sum_{i=2}^{d} x_i^2 \quad (d = 10). \tag{11}$$

The double well potential we consider is given by:

$$V(\mathbf{x}) = \left(x_1^2 - 1\right)^2 + 0.3 \sum_{i=2}^{d} x_i^2 \quad (d = 10). \tag{11}$$

The regions A and B are defined as

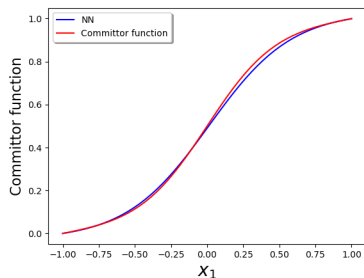$$A = \left\{ x \in \mathbb{R}^d \mid x_1 \leq -1 \right\}, \quad B = \left\{ x \in \mathbb{R}^d \mid x_1 \geq 1 \right\}.$$

Numerical experiment I – Double well potential

The double well potential we consider is given by:

$$V(\mathbf{x}) = \left(x_1^2 - 1\right)^2 + 0.3 \sum_{i=2}^{d} x_i^2 \quad (d = 10). \tag{11}$$

The regions A and B are defined as

$$A = \left\{ x \in \mathbb{R}^d \mid x_1 \leq -1 \right\}, \quad B = \left\{ x \in \mathbb{R}^d \mid x_1 \geq 1 \right\}.$$

T	E	No. training samples	Batch size	No. testing samples
0.5	0.014	1.5×10^5	1000	4.0×10^5
0.2	0.011	8.0×10^5	1000	8.0×10^5

Table 1: Results for the double-well potential problem.

The double well potential we consider is given by:

$$V(\mathbf{x}) = \left(x_1^2 - 1\right)^2 + 0.3 \sum_{i=2}^{d} x_i^2 \quad (d = 10). \tag{11}$$

The regions A and B are defined as

$$A = \left\{x \in \mathbb{R}^d \mid x_1 \leq -1\right\}, \quad B = \left\{x \in \mathbb{R}^d \mid x_1 \geq 1\right\}.$$

T	E	No. training samples	Batch size	No. testing samples
0.5	0.014	1.5×10^5	1000	4.0×10^5
0.2	0.011	8.0×10^5	1000	8.0×10^5

Table 1: Results for the double-well potential problem.

Here the error E is defined by $E = \frac{\|q_\theta - q^*\|_{L_\rho^2(\Omega \setminus A \cup B)}}{\|q^*\|_{L_\rho^2(\Omega \setminus A \cup B)}}$, where q^* is the ground truth.

Numerical experiment I – Double well potential

Figure 1: The committor function for the double-well potential along x_1 dimension when $T = 0.5$ for an arbitrarily chosen $(x_2, \ldots, x_d)$.
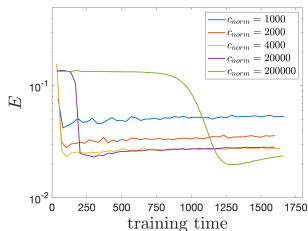
Numerical experiment I – Double well potential
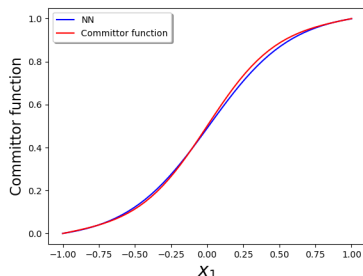
Figure 1: The committor function for the double-well potential along x_1 dimension when $T = 0.5$ for an arbitrarily chosen $(x_2, \ldots, x_d)$.

The final error is not sensitive to the parameter δ.
If $\delta = 0.01, 0.03, 0.05$ are chosen instead, the corresponding final errors are $E = 0.013, 0.013, 0.013$, respectively.

Sensitivity of the penalty coefficient
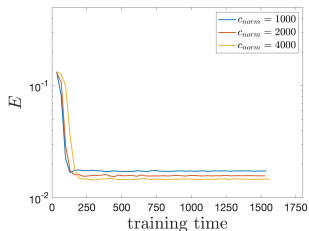
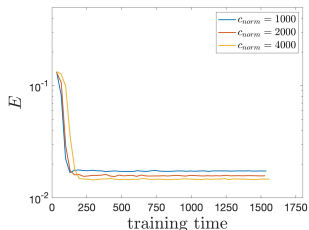(a) Training loss of the proposed method

(b) Training loss of the method in [2]

Figure 2: Comparison of the training process of the proposed method and [2]'s method. Here c_{norm} stands for the normalized penalty coefficient.

Sensitivity of the penalty coefficient
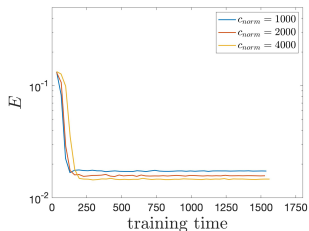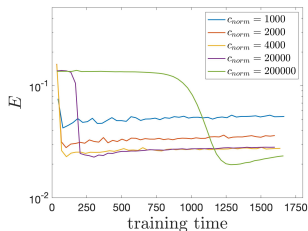
(a) Training loss of the proposed method
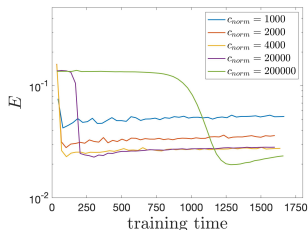
(b) Training loss of the method in [2]

Figure 2: Comparison of the training process of the proposed method and [2]'s method. Here c_{norm} stands for the normalized penalty coefficient.

- In Fig. 2a, the approximate solution converges quickly and the final relative error is rather small, regardless of the choice of penalty coefficients.

Sensitivity of the penalty coefficient

(a) Training loss of the proposed method

(b) Training loss of the method in [2]

Figure 2: Comparison of the training process of the proposed method and [2]'s method. Here c_{norm} stands for the normalized penalty coefficient.

- In Fig. 2a, the approximate solution converges quickly and the final relative error is rather small, regardless of the choice of penalty coefficients.
- From Fig. 2b, we can see that the training process and final result is sensitive to c_{norm}.

Numerical experiment II – Rugged-Mueller potential

The rugged-Mueller potential is given by:

$$V(\mathbf{x}) = \tilde{V}(x_1, x_2) + \frac{1}{2\sigma^2} \sum_{i=3}^{d} x_i^2,$$

where $\tilde{V}$ is the 2-dimensional rugged Mueller potential. The domain of interest Ω of this example is $[-1.5, 1] \times [-0.5, 2] \times \mathbb{R}^{d-2}$ and the regions A and B are the following two cylinders:

$$A = \left\{ \mathbf{x} \in \mathbb{R}^d \mid \sqrt{(x_1 + 0.57)^2 + (x_2 - 1.43)^2} \leq 0.3 \right\},$$

$$B = \left\{ \mathbf{x} \in \mathbb{R}^d \mid \sqrt{(x_1 - 0.56)^2 + (x_2 - 0.044)^2} \leq 0.3 \right\}.$$
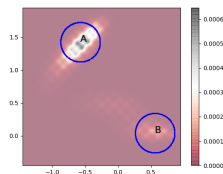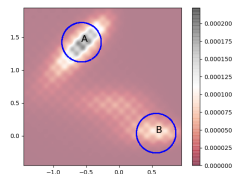
Numerical experiment II – Rugged-Mueller potential

The rugged-Mueller potential is given by:

$$V(\mathbf{x}) = \tilde{V}(x_1, x_2) + \frac{1}{2\sigma^2} \sum_{i=3}^{d} x_i^2,$$

where $\tilde{V}$ is the 2-dimensional rugged Mueller potential. The domain of interest Ω of this example is $[-1.5, 1] \times [-0.5, 2] \times \mathbb{R}^{d-2}$ and the regions A and B are the following two cylinders:

$$A = \left\{ \mathbf{x} \in \mathbb{R}^d \mid \sqrt{(x_1 + 0.57)^2 + (x_2 - 1.43)^2} \leq 0.3 \right\},$$

$$B = \left\{ \mathbf{x} \in \mathbb{R}^d \mid \sqrt{(x_1 - 0.56)^2 + (x_2 - 0.044)^2} \leq 0.3 \right\}.$$
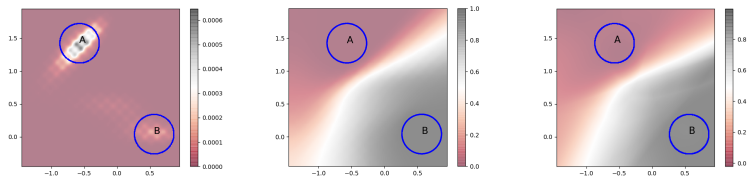

(a) $T = 22$ (b) $T = 40$

Figure 3: The equilibrium distributions for the rugged-Mueller potential.

Numerical experiment II – Rugged-Mueller potential

The results are summarized in Table 2.

(T, σ)	E	No. training samples	Batch size	No. testing samples
(22, 0.05)	0.024	6.0×10^5	5000	1.0×10^6
(40, 0.05)	0.023	6.0×10^5	5000	1.0×10^6

Table 2: Results for the rugged-Muller potential problem.

(a) equilibrium distribution   (b) $T = 22$ committor function (c) $T = 22$ NN approximation

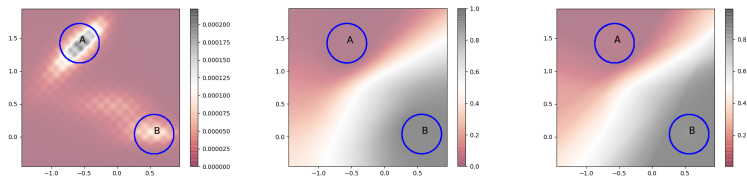Figure 4: Comparisons between the NN represented committor functions and the ground truth when $T = 22$.

Numerical experiment II – Rugged-Mueller potential

The results are summarized in Table 2.

(T, σ)	E	No. training samples	Batch size	No. testing samples
(22, 0.05)	0.024	6.0×10^5	5000	1.0×10^6
(40, 0.05)	0.023	6.0×10^5	5000	1.0×10^6

Table 2: Results for the rugged-Muller potential problem.

(a) equilibrium distribution (b) $T = 40$ committor function (c) $T = 40$ NN approximation

Figure 4: Comparisons between the NN represented committor functions and the ground truth when $T = 40$.

Numerical experiment III – Ginzburg-Landau potential

The Ginzburg-Landau energy in one dimension is defined as:

$$\tilde{V}[u] = \int_0^1 \frac{\lambda}{2} u_x^2 + \frac{1}{4\lambda}(1-u^2)^2 dx, \tag{12}$$

where λ is a small positive parameter and u is a sufficiently smooth function on $[0,1]$ with boundary conditions $u(0) = u(1) = 0$.

Numerical experiment III – Ginzburg-Landau potential

The Ginzburg-Landau energy in one dimension is defined as:

$$\tilde{V}[u] = \int_0^1 \frac{\lambda}{2} u_x^2 + \frac{1}{4\lambda}(1 - u^2)^2 dx, \tag{12}$$

where λ is a small positive parameter and u is a sufficiently smooth function on $[0, 1]$ with boundary conditions $u(0) = u(1) = 0$.

$u(x)$ can be uniformly discretized by $U = (U_1, \cdots, U_d)$ defined on a uniform grid on $[0, 1]$ with the boundary conditions $U_0 = U_{d+1} = 0$. Then the continuous Ginzburg-Landau energy is approximated by:

$$V(U) := \tilde{V}_h[U] = \sum_{i=1}^{d+1} \frac{\lambda}{2} \left(\frac{U_i - U_{i-1}}{h} \right)^2 + \frac{1}{4\lambda}(1 - U_i^2)^2, \tag{13}$$

where the grid size $h = 1/(d + 1)$.

Numerical experiment III – Ginzburg-Landau potential

The Ginzburg-Landau energy in one dimension is defined as:

$$\tilde{V}[u] = \int_0^1 \frac{\lambda}{2} u_x^2 + \frac{1}{4\lambda}(1 - u^2)^2 dx, \tag{14}$$

where λ is a small positive parameter and u is a sufficiently smooth function on $[0,1]$ with boundary conditions $u(0) = u(1) = 0$.

Numerical experiment III – Ginzburg-Landau potential

The Ginzburg-Landau energy in one dimension is defined as:

$$\tilde{V}[u] = \int_0^1 \frac{\lambda}{2} u_x^2 + \frac{1}{4\lambda}(1 - u^2)^2 dx, \tag{14}$$

where λ is a small positive parameter and u is a sufficiently smooth function on $[0, 1]$ with boundary conditions $u(0) = u(1) = 0$.

$u(x)$ can be uniformly discretized by $U = (U_1, \cdots, U_d)$ defined on a uniform grid on $[0, 1]$ with the boundary conditions $U_0 = U_{d+1} = 0$. Then the continuous Ginzburg-Landau energy is approximated by:

$$V(U) := \tilde{V}_h[U] = \sum_{i=1}^{d+1} \frac{\lambda}{2} \left(\frac{U_i - U_{i-1}}{h} \right)^2 + \frac{1}{4\lambda}(1 - U_i^2)^2, \tag{15}$$

where the grid size $h = 1/(d + 1)$. In this experiment we use $h = 1/50$ and the dimension $d = 49$.

$V(U)$ has two local minima $u_\pm(\cdot)$ as shown in Fig. 5. The regions A and B are taken as the spheres $\{U : ||U - u_\pm|| \leq r\}$ with $r = 3$.

Numerical experiment III – Ginzburg-Landau potential

$V(U)$ has two local minima $u_\pm(\cdot)$ as shown in Fig. 5. The regions A and B are taken as the spheres $\{U : ||U - u_\pm|| \le r\}$ with $r = 3$.
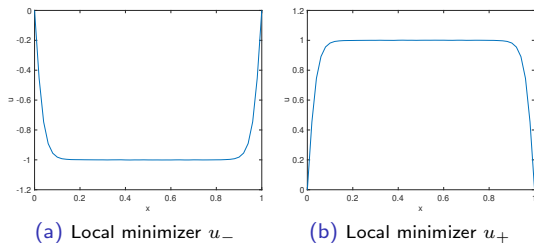

(a) Local minimizer u_- (b) Local minimizer u_+

Figure 5: Two local minima of the energy (15) with $\lambda = 0.03$. (a): u_-, (b): u_+.

Verification of the result

$$\Gamma_{\frac{1}{2},\epsilon} = \{U : |q_\theta(U) - \frac{1}{2}| < \epsilon\}$$

We verify the numerical result in the following way:

- Identify m states $\{\tilde{\mathbf{x}}_j\}_{j=1}^m$ on $\Gamma_{\frac{1}{2},\epsilon}$,

Verification of the result

$$\Gamma_{\frac{1}{2},\epsilon} = \{U : |q_\theta(U) - \frac{1}{2}| < \epsilon\}$$

We verify the numerical result in the following way:

- Identify m states $\{\tilde{\mathbf{x}}_j\}_{j=1}^m$ on $\Gamma_{\frac{1}{2},\epsilon}$,
- for each $\tilde{\mathbf{x}}_j$, generate N trajectories according to (1),

Numerical experiment III – Ginzburg-Landau potential

Verification of the result

$$\Gamma_{\frac{1}{2}, \epsilon} = \{U : |q_\theta(U) - \frac{1}{2}| < \epsilon\}$$

We verify the numerical result in the following way:

- Identify m states $\{\tilde{\mathbf{x}}_j\}_{j=1}^m$ on $\Gamma_{\frac{1}{2}, \epsilon}$,
- for each $\tilde{\mathbf{x}}_j$, generate N trajectories according to (1),
- denote the number of trajectories that start from $\tilde{\mathbf{x}}_j$ and reach B before A as n_j,

Verification of the result

$$\Gamma_{\frac{1}{2},\epsilon} = \{U : |q_\theta(U) - \frac{1}{2}| < \epsilon\}$$

We verify the numerical result in the following way:

- Identify m states $\{\tilde{\mathbf{x}}_j\}_{j=1}^m$ on $\Gamma_{\frac{1}{2},\epsilon}$,
- for each $\tilde{\mathbf{x}}_j$, generate N trajectories according to (1),
- denote the number of trajectories that start from $\tilde{\mathbf{x}}_j$ and reach B before A as n_j,
- compare the distribution of n/N with $\mathcal{N}(\frac{1}{2}, (4N)^{-1})$, i.e. the normal distribution with mean $\frac{1}{2}$ and variance $(4N)^{-1}$.

Numerical experiment III – Ginzburg-Landau potential

The comparison of the distribution of n/N with $\mathcal{N}(\frac{1}{2}, (4N)^{-1})$ is given in the following figures. In the actual experiment with $\epsilon = 0.01$, $m = 120$, and $N = 100$, the resulting statistics contain n_j/N for $j = 1, 2, \ldots, 120$.
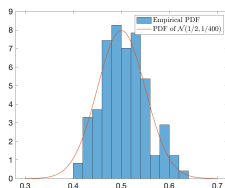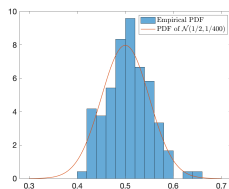

(a) The empirical PDF versus the PDF of $\mathcal{N}(\frac{1}{2}, 1/400)$.
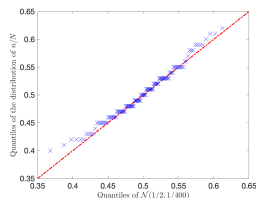
(b) Q–Q plot of $\{n_j/N\}_{j=1}^{120}$ versus $\mathcal{N}(\frac{1}{2}, 1/400)$.

Numerical experiment III – Ginzburg-Landau potential

The comparison of the distribution of n/N with $\mathcal{N}(\frac{1}{2}, (4N)^{-1})$ is given in the following figures. In the actual experiment with $\epsilon = 0.01$, $m = 120$, and $N = 100$, the resulting statistics contain n_j/N for $j = 1, 2, \ldots, 120$.

(c) The empirical PDF versus the PDF of $\mathcal{N}(\frac{1}{2}, 1/400)$.
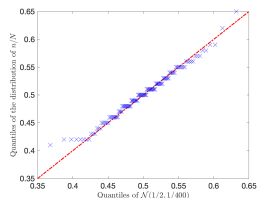
(d) Q–Q plot of $\{n_j/N\}_{j=1}^{120}$ versus $\mathcal{N}(\frac{1}{2}, 1/400)$

Conclusion

Summary:

- We show that the committor function satisfies an integral equation based on the semigroup of the Fokker-Planck operator.

Summary:

- We show that the committor function satisfies an integral equation based on the semigroup of the Fokker-Planck operator.

- The explicit gradient is removed and the boundary conditions are handled naturally.

Conclusion

Summary:

- We show that the committor function satisfies an integral equation based on the semigroup of the Fokker-Planck operator.

- The explicit gradient is removed and the boundary conditions are handled naturally.

- The integrals in the variational form is approximated via sampling, and the committor function is solved for using NN parameterization and SGD.

Summary:

- We show that the committor function satisfies an integral equation based on the semigroup of the Fokker-Planck operator.

- The explicit gradient is removed and the boundary conditions are handled naturally.

- The integrals in the variational form is approximated via sampling, and the committor function is solved for using NN parameterization and SGD.

- The convergence of the training process is guaranteed in the lazy training regime.

Conclusion

Summary:

- We show that the committor function satisfies an integral equation based on the semigroup of the Fokker-Planck operator.

- The explicit gradient is removed and the boundary conditions are handled naturally.

- The integrals in the variational form is approximated via sampling, and the committor function is solved for using NN parameterization and SGD.

- The convergence of the training process is guaranteed in the lazy training regime.

- The resulting algorithm is shown to be faster and less sensitive to the penalty parameter then the previous work.

Conclusion

Future work:

- Integrate importance sampling techniques in the sampling process.
- Adopt higher order integration schemes instead of Euler-Maruyama.
- Apply the proposed method to other high-dimensional PDEs and eigenvalue problems that possess probabilistic interpretations (See our more recent work [5]).

[1] E. Vanden-Eijnden and E. Weinan, "Modeling rare transition events,",

[2] Y. Khoo, J. Lu, and L. Ying, "Solving for high-dimensional committor functions using artificial neural networks," *Research in the Mathematical Sciences*, vol. 6, no. 1, p. 1, 2019.

[3] H. Li, Y. Khoo, Y. Ren, and L. Ying, "A semigroup method for high dimensional committor functions based on neural network," *arXiv preprint arXiv:2012.06727, accepted by the Mathematical and Scientific Machine Learning (MSML 2021) conference*, 2020.

[4] E. Dynkin, "Markov processes.," *New York: Academic Press Inc.*, 1965.

[5] H. Li and L. Ying, "A semigroup method for high dimensional elliptic pdes and eigenvalue problems based on neural networks," *arXiv preprint arXiv:2105.03480*, 2021.

Thank you for your attention!